数据脱敏4 | 法律可以量化评价数据脱敏的效果吗？

合规科技系列文章 **Law-Tech Series**

高速发展的时代背景下，一方面行业分工在层层细化，一方面跨学科的交叉研究又越来越不可或缺。科技与法律表面上是两个相去甚远的专业领域，但就数据治理与隐私保护而言，只有跨界互通才可能找到最佳的解决方案。

“合规科技专题文章”旨在兼顾科技与法律的双重视角，深度解读数据技术的逻辑原理与数据合规的法律要求，从而促进技术人与法律人的双向理解，探讨数据利用与个人权益协调发展的可行方案。

“大数据”已然从热词变成日常，而数据在释放无限潜力的同时，也引发了隐私泄露的巨大隐患。从若干年前科技公司野蛮生长，到近年来数据立法接踵而至，信息社会正在两极之间寻求平衡。数据脱敏提供了这样一种可能性——通过降低数据与主体之间的关联，可以同时保留较高的隐私保护程度和较大的数据利用价值。

“数据脱敏”专题文章将梳理匿名化、去标识化、假名化等一系列相关概念，分析中国、欧盟、美国等法域对不同概念的法律评价，介绍数据脱敏的技术方案与隐私模型，探讨各个业务场景下的行业实践案例与法律落地方案，以推动数据利用和隐私保护的平衡发展。

“数据脱敏”专题往期文章链接

- 数据脱敏1 | “数据脱敏”是一个法律概念或技术概念吗？
- 数据脱敏2 | 不同法域下匿名化、去标识化、假名化的含义一致吗？
- 数据脱敏3 | 脱敏技术与法律效果评价可以机械对应的吗？
- 数据脱敏4 | 法律可以量化评价数据脱敏的效果吗？

上期回顾：数据脱敏可以采用统计、密码、抑制、假名化、泛化、随机化、数据合成等技术。法律对脱敏技术的评价并非机械对应，而是考量特定的技术方案、实施强度和应用场景，具体评价其实现的不可识别的程度。

那么，法律上如何衡量脱敏的效果，即不可识别的程度呢？本文将介绍传统的定性标准（如第三人标准、安全港标准）和专门的定量标准（如K-匿名、差分隐私模型），并探讨通过数据分析进一步精细化的量化评价路径。

一. 评价脱敏效果的定性标准

法律上的标准往往不是精确的数字，而是定性的描述。就脱敏效果的标准而言，各国法上都有第三人标准，同时，也有立法试图列举应当被“脱”的数据项。

1. 第三人标准

第三人标准是常见的立法例，各国法上的主要差异在于第三方的性质和范围。

美国《健康保险流通与责任法案》（HIPAA）对健康数据的去标识化提出了专家测定标准（**Expert Determination**），即经过具备统计知识与科学方法的专家的测定，考虑到数据接收者合理可得的数据和合理可用的技术，从脱敏后的健康信息中识别出个人的风险非常小。

英国信息专员办公室（Information Commissioners Office）在《匿名化：管理数据保护风险的实践准则》中提出了的有动机的入侵者测试（**Motivated Intruder Test**），即对于并非内幕人士或专业黑客的一般第三人而言，通过公开检索、询问、调查等方式，匿名信息能否被重识别。

我国国家标准《个人信息去标识化指南》则分别提出了针对外部一般人员和内部违规人员的入侵者测试。

2. 安全港标准

美国HIPPA就健康信息的去标识化提出了安全港标准（**Safe Harbor**），指去除18项标识符，并且数据处理者不认为处理后的信息能够单独或结合地识别个人。这18项标识符包括姓名、小于州的地理信息、小于年的日期信息、电话号码、传真号码、电子邮箱地址、社会保险号、病历号、健康计划受益人号、银行账号、证书号、车辆识别号、设备识别符、URL地址、IP地址、生物识别符、正面照片、其他唯一识别符号。

与之类似的是，上海市卫生健康委员会在新冠疫情流调报告中去除了确诊病例的姓名、性别、年龄等标识符，仅公开时间、区域和场所等信息，从而保护了病人的隐私。

但是，HIPPA的安全港标准也受到批评——18项标识符的列举并不周延，并且删除标识符并不意味着去标识化，尤其对于较小的样本量或罕见的属性值，个人仍然可能被识别。例如，1000人中的Rh阴性血很可能指向唯一的个人。

二. 评价脱敏效果的定量标准

我国的《个人信息去标识化指南》、欧盟的《关于匿名化技术的意见》、国际标准化组织的《隐私增强数据去标识化术语和技术分类》（ISO/IEC 20889）都介绍了K-匿名模型、差分隐私模型，运用数学的方法为数据脱敏和隐私保护提供了定量的评价标准。

1. K-匿名模型

哈佛大学教授Latanaya Sweeney研究指出，结合出生日期、性别、邮政编码这三项属性可以识别出87%的美国人，因此她于1998年提出了K-匿名。**K-**匿名（***K-***

anonymity）是指，对某一标识符进行一定程度的泛化，使得对于任一属性值，至少有K个数据主体共享同一属性值。

如下图，假设有一个可供公开查询的数据库，包含出生日期、住址、患病情况这三项属性。一个攻击者已知其攻击目标张三的出生日期（1950年2月1日）、住址（北京市朝阳区建外街道幸福小区），并试图获取张三的患病情况。

泛化处理	出生日期	住址	符合前两项的人数	患病情况	攻击结论
原始数据	1950年2月1日	北京市朝阳区建外街道幸福小区	1 （K=1）	1人高血压	张三一定患有高血压
泛化	1950年2月	北京市朝阳区建外街道	10 （K=10, L=1）	10人心脏病	张三一定患有高血压
再泛化	1950年	北京市朝阳区	200 （L=2）	10人心脏病 190人高血压	张三大概率患有高血压

在原始数据库，攻击者通过出生日期和住址就能锁定张三，从而暴露其患有高血压。如果将出生日期泛化至出生年月、将小区泛化至街道，则与张三共享这两项属性的人增加至10人（K=10），因此攻击者无法从等价类中识别出张三。但是，如果攻击所针对的敏感属性（患病情况）的属性值差异很小，则K-匿名的效果有限。例如，居住在建外街道、1950年2月出生的10人都患有高血压，则攻击者可以确定其张三也患有高血压。

L-多样性（***L*-diversity**）在K-匿名的基础上，要求每一等价类在每一敏感属性上至少有L个不同的属性值。例如，当进一步泛化至居住在北京市朝阳区、1950年出生的人，这200人患有高血压或心脏病（L=2，指两种疾病），则攻击者无法确定张三患有哪种病。但是，如果敏感属性值的分布不均，则L-多样性的效果有限。例如，200人中有10个心脏病，190个高血压，则攻击者可以推理出张三大概率也患有高血压。

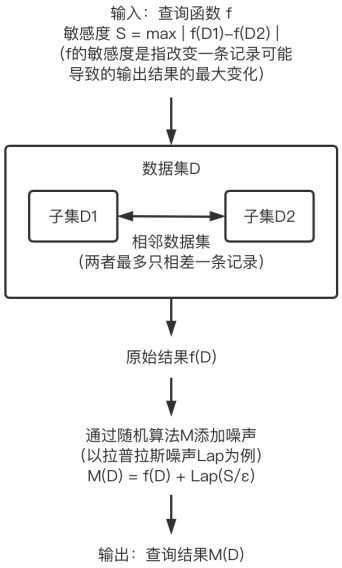
T-接近性（***T*-closeness**）在L-多样性的基础上，要求敏感属性在任何等价类中的分布与其在整个数据集中的分布相近，两者差距小于阈值T。但对于发病概率与年龄强相关的疾病，比较难通过泛化出生日期来实现T-接近性。例如，老年人更容易得高血压，则在张三的等价类中，高血压的患者可能远多于整个数据集平均情况。

2. 差分隐私模型

差分隐私（**differential privacy**）是指，通过向数据集中添加随机噪声，使得任意个人的记录对该数据集或其子集的统计特性影响极小。这种噪声（如拉普拉斯噪声）是经过精心设计、符合概率分布的，从而使攻击者无法从数据集的查询结果及其组合中筛选出某一个人，但这不可避免地导致数据在一定程度上失真。

例如，一个社区中高血压的患病人数为40人，当新搬来一个住户后，如果患病人数变为41人，则可以判断新住户患有高血压，导致其隐私泄露（统计数据并不一定是匿名的）。添加噪声之后，假设患病人数仍为40人，但40是一个有噪声的、非确切的数字，因此无法判断新住户的患病情况。

差分隐私模型为隐私保护程度提供了严格的数学定义。如下图，基于给定的隐私预算ε（指愿意忍受的隐私风险），对于数据集中任意两个相邻数据集D1、D2，当一个查询函数f的敏感度为S，如果某一随机算法M满足概率公式$Pr[M(D1)] \leq \exp(\varepsilon) \times Pr[M(D2)]$，则该算法M实现了“ε-差分隐私”。



三. 进一步的量化标准

上述的定性标准和定量标准都为脱敏效果的衡量提供了参考，但是，它们都没有彻底回答衡量标准的问题。

第三人标准聚焦于重识别的主体，各国分别界定了第三人的性质和范围，如专家、内幕人员、外部的一般第三人等。但是，第三人标准并没有明确第三人在测试数据集时应当使用的方法和标准，因此，企业难以自查和判断数据脱敏的效果。

安全港标准试图列举出若干的标识符，引导企业删除这些敏感的属性值，这或许可以在特定行业内提供脱敏的最低门槛。但是，数据不是孤立的，数据之间是存在联系的，即使去除了标识符，若干的非敏感属性值相结合，也可能推导出数据主体的身份。

K-匿名模型和差分隐私模型突破了传统的定性标准，对脱敏的程度进行了数学上的定义，为效果的度量提供了量化的工具。但是，各国法上只给出了K、L、T、ε等参数，却并未对其进行赋值。K＞10就是去标识化吗？K＞10000就是匿名化吗？数据脱敏真的有明确的量化标准吗？

这些数字或许永远不会有标准答案，但实践中不妨借助数据的力量，模拟出一个相对科学的衡量标准。例如，就同一行业、相似的业务场景抽样100家企业及其数据集，并设定统一的数据可用性需求和隐私保护要求，再由各个企业分别对其数据集进行脱敏处理。对于脱敏后的数据集，通过K-匿名或差分隐私模型进行验证，计算出各个数据集的K/L/T/ε值，并对这100组数值进行排列与分析，从而大致推算出当前实践中可以接受的K/L/T/ε标准。

本期小结与下期预告：对于脱敏效果的衡量，既有定性的第三人标准、安全港标准，也有定量的K-匿名模型、差分隐私模型。目前，不可识别的程度并没有精确的度量方法，但可以通过特定场景下的实证研究，进一步探索量化的标准。那么，既然不可识别的程度是渐进的，数据脱敏的法律效果也是渐进的吗？下期文章将为您分析现行法下假名化、去标识化、匿名化的法律地位。